

*Etude sur l'offre numérique
illégal des livres français
sur Internet en 2010*
2^e année

ENFOOKZ

2

ElabZ

Mars 2011

Auteur

Mathias Daval

(Edysseus Consulting)
pour le MOTif

le MOT if

Cécile Moscovitz
Responsable

des études

6, villa Marcel-Lods
Passage de l'Atlas

Paris 19^e

Métro Belleville

01 53 38 60 61

contact@lemotif.fr

www.lemotif.fr

le MOT if

est un organisme associé
de la Région Ile-de-France.

 **ile de France**

Table des matières

Introduction.....	3
Champ de l'étude.....	4
Mise à jour des principales et récentes études évoquant le téléchargement illégal des livres.....	5
Evolution du piratage de livres sur Internet entre octobre 2009 et janvier 2010.....	6
Méthodologie de l'étude	10
Les circuits de téléchargement illégal.....	10
Une évaluation globale de l'offre.....	16
Echantillonnage	17
Catégories éditoriales.....	18
Résultats	19
Analyse des fichiers.....	19
Analyse des ouvrages	23
Synthèse des résultats.....	33
Lexique	35

Introduction

L'étude « EbookZ 1 », publiée en octobre 2009, a été la première en France à analyser spécifiquement l'**offre illégale** des livres sur Internet, et à répondre notamment aux questions suivantes :

- Quelle est la nature de cette offre ?
- Quels sont les livres et les éditeurs les plus concernés ?
- Par quels circuits et comment sont piratés ces ouvrages ?
- De quelles quantités parle-t-on ?

Afin d'étudier dans la durée le piratage de livres sur le net, le MOTif a créé un outil pour mesurer le phénomène : ElabZ, qui a aussi plus largement vocation à décrypter les usages de la consommation de livres numériques.

Dans le cadre d'ElabZ, la présente étude est donc une mise à jour d'« EbookZ 1 », 18 mois après : après le lancement du Kindle et de l'iPad en France, et la démultiplication rapide des plates-formes de distribution (cf. l'étude « Tableau de bord n°1 – Offre de livres numériques », octobre 2010), il semblait pertinent d'étudier dans quelle mesure l'offre pirate avait évolué et conservé ou non ses caractéristiques telles que nous les avons définies dans notre première étude.

Dans un but comparatif, nous avons donc, sauf exception mentionnée comme telle, utilisé la **même méthodologie** que dans « EbookZ 1 » (cf. rappel p. 10).

Nous avons mis à jour la synthèse des différentes études publiées sur le piratage des livres (p. 5).

Nous avons également ajouté 2 focus spécifiques :

- L'un sur **la bande dessinée**, après une collecte exhaustive des fichiers disponibles sur l'un des principaux agrégateurs ed2K (p. 27).
- L'autre sur **la littérature professionnelle** (p. 30).

NB : Les termes de l'étude suivis d'un astérisque* sont définis dans le lexique page 35.

Champ de l'étude

Contenus — Rappel des principes de l'étude « EbookZ »

Le sujet de l'étude est le « piratage numérique du livre » et non le « piratage du livre numérique » : il n'est pas restreint aux seuls titres disponibles en offre numérique **légale**, mais bien à l'ensemble des textes piratés, dont l'essentiel provient de livres papier scannés et diffusés en ligne. Lorsque nous parlerons de « livres piratés », il s'agira donc de l'ensemble des caractéristiques de **l'offre numérique illégale**, qu'elle provienne du scannage de livres imprimés ou du *cracking* (« craquage ») de livres déjà existants au format numérique.

La notion de « livre numérique » ou **ebook** suppose une redéfinition juridique du mot « livre », jusqu'alors traditionnellement réservée aux imprimés non périodiques (« *Un livre est un ensemble imprimé, illustré ou non, publié sous un titre ayant pour objet la reproduction d'une œuvre de l'esprit d'un ou plusieurs auteurs en vue de l'enseignement, de la diffusion de la pensée et de la culture* », telle est la définition fiscale du livre), à l'exclusion donc de la presse¹.

Or l'usage courant du mot « **ebook** » en matière de piratage, que l'on retrouve aussi sous la forme « **ebookz** » ou « **bookwarez** », recouvre les livres mais aussi les périodiques. En pratique, ces derniers constituent même 70 à 80 % du nombre de fichiers disponibles en téléchargement illégal². Nous les avons systématiquement exclus des fichiers analysés.

L'étude ne concerne que les **contenus payants et sous droits**, donc tombant sous le coup de la législation française sur le droit d'auteur. La mise à disposition gratuite de contenus numériques issus du domaine public, mais qui peuvent constituer une partie non négligeable du chiffre d'affaires de certains éditeurs, est un enjeu qui déborde le cadre de notre étude et qui concerne une quantité infime des échanges en matière de peer to peer*.

Enfin l'étude ne tiendra pas compte du piratage des traductions d'œuvres françaises. Elle ne concerne que les **livres publiés par des éditeurs français, en français** (traduites ou non).

Fichiers analysés

Nous avons défini des échantillons ou ensembles d'ouvrages disponibles au téléchargement illégal, divisé en 2 catégories : les **livres** et les **bandes dessinées (BD)**. Nous avons exclu, pour cette mise à jour, les livres audio.

Ces ensembles ont été constitués par les résultats de recherches croisées sur différents circuits de diffusion et affiniées en excluant les fichiers corrompus ou ne correspondant pas à des livres sous droits.

¹ Nous avons également exclu de l'étude les partitions musicales, qui font l'objet d'un piratage massif et spécifique.

² D'après notre observation des rubriques « ebooks » des principaux *trackers** et sites de référence en peer to peer*.

Mise à jour des principales et récentes études évoquant le téléchargement illégal des livres (novembre 2009 à janvier 2010)

« Pour un livre numérique créateur de valeurs » (rapport Albanel)

Remis au Premier ministre en avril 2010³, le rapport commandité à Christine Albanel en novembre 2009 comporte 3 axes dont le second « concerne l'adaptation du secteur de l'édition à la lutte contre le piratage de ses contenus ».

Notre étude « EbookZ 1 » d'octobre 2009 est la seule source mentionnée à propos du piratage, dont le rapport reprend les principales conclusions chiffrées sur l'offre disponible, et ajoute : « C'est d'ailleurs pourquoi le meilleur rempart contre ce fléau est le développement d'un modèle offrant la plus grande richesse de contenus dans des conditions d'utilisation les plus simples possibles, ce qui correspond aux attentes d'un lecteur, qui est aussi un usager quotidien de l'internet, avec les réflexes liés à sa pratique du réseau. »

Le rapport conclut sous forme de conseil juridique aux éditeurs : « Dès lors, si les éditeurs se regroupaient et adhéreraient collectivement à l'ALPA, un contrôle du téléchargement illégal ou de tout mode illégal de mise à disposition de fichiers de livres numériques pourrait être mis en œuvre. »

Les études d'Attributor

Attributor est une société américaine spécialisée dans la surveillance des réseaux pirates pour le compte des ayants-droit⁴.

En janvier 2010, elle a publié les résultats d'une étude portant sur l'estimation des volumes de téléchargement d'ebooks pirates aux Etats-Unis. A partir d'un panel de 913 titres représentatifs de la production éditoriale du pays, elle a suivi leur mise à disposition illégale sur Internet pendant 3 mois (octobre-décembre 2009) :

- 9 millions de téléchargements, soit environ 10 000 téléchargements par jour.
- Une perte pour l'industrie du livre estimée à environ 2,75-3 Mds de \$, soit 10 % du chiffre d'affaires global de l'édition aux Etats-Unis.

Dans une deuxième étude parue en octobre 2010, controversée⁵, Attributor estime de 1,5 à 3 millions le nombre de téléchargeurs quotidiens d'ebooks illégaux. Attributor en conclut que nous vivons actuellement le « moment Napster » du livre, c'est-à-dire une phase où l'offre illégale concurrence sévèrement l'offre légale. Mais que celle-ci peut aussi être attractive : l'étude montre que près de 20 % des internautes ayant abouti à un lien de téléchargement Amazon ont cliqué sur ce lien menant vers la possibilité d'achat légal du livre.

³ <http://www.gouvernement.fr/presse/le-premier-ministre-a-recu-le-rapport-de-christine-albanel-sur-le-livre-numerique>

⁴ http://www.attributor.com/solutions/book_monitor.php

⁵ <http://go-to-hellman.blogspot.com/2010/10/attributor-ebook-piracy-numbers-dont.html>

Publié en juin 2010, l'ouvrage de Joël Faucilhon est une synthèse augmentée de différents billets parus dans son blog Lektı-Ecriture⁶. Il porte sur l'Internet comme utopie et notamment sur le rôle des communautés de pirates comme réappropriation du web, qu'on « aurait tort de considérer simplement comme de l'activisme dédié à la destruction des économies de la culture ». Il reprend les analyses d'Hakim Bey (cf. « EbookZ 1 »), en montrant que certaines communautés pirates privées comme Bitme ou Elbitz pour les livres fonctionnent sur le modèle de TAZ (« zones autonomes temporaires ») propres au web⁷. Il y est par ailleurs question de l'inefficacité des DRM dans la lutte contre le piratage.

L'étude de l'Hadopi (janvier 2011)

La Haute Autorité pour la diffusion des œuvres et la protection des droits sur Internet a publié fin janvier les résultats d'une première étude, dans le cadre d'un baromètre sur les pratiques de téléchargement des français. Cette étude porte sur l'ensemble des biens culturels et pas seulement les livres : 49 % des internautes⁸ déclarent « avoir consommé au moins une fois sur Internet des produits culturels de manière illicite », dont 29 % depuis moins de 6 mois, avec un découpage par tranche d'âge bien marqué : 70 % des 15-24 ans, 55 % des 25-39 et 32 % des 40 et plus.

Chez les internautes déclarant un usage illicite, les livres arrivent en 3^e position des téléchargements légaux (après la musique et les photos), mais seulement en 7^e et dernière position des téléchargements illégaux.

L'étude souligne aussi que « 75% des internautes dépensent en moyenne 36 euros par mois en biens culturels (achats par correspondance inclus). Les internautes déclarant un usage illicite ont une dépense moyenne supérieure ».

Evolution du piratage de livres sur Internet entre octobre 2009 et janvier 2010

Il faut noter deux apparitions dans l'univers de l'ebook pirate francophone pendant cette période :

1°) Une nouvelle team* consacrée aux ebooks, avec une double intention quantitative et qualitative.

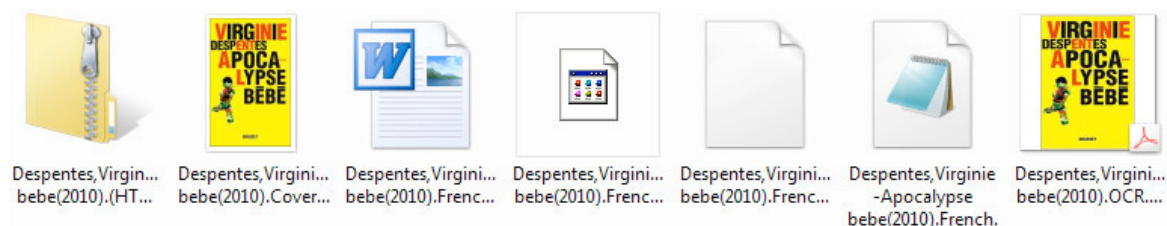
L'équipe pirate la plus renommée en France aujourd'hui est aussi l'une des plus récentes. Elle est composée d'une douzaine de membres permanents auxquels s'ajoutent des intervenants occasionnels. Elle dispose d'un forum de discussion, mais elle possède aussi un compte Twitter qu'elle utilise pour avertir des nouvelles « *releases** » ou sorties de publications illégales, ainsi qu'une syndication OPDS (protocole combinant les caractéristiques des flux RSS et du peer to peer, notamment utilisable sur les appareils mobiles).

⁶ <http://www.lekti-ecriture.com/editeurs/Reveurs-marchands-et-pirates.html>

⁷ Les TAZ sont des sortes d'enclaves libres, régies par leurs propres lois, et réunissant des participants partageant des valeurs communes.

⁸ Échantillon de 2 687 internautes âgés de 15 ans et plus.

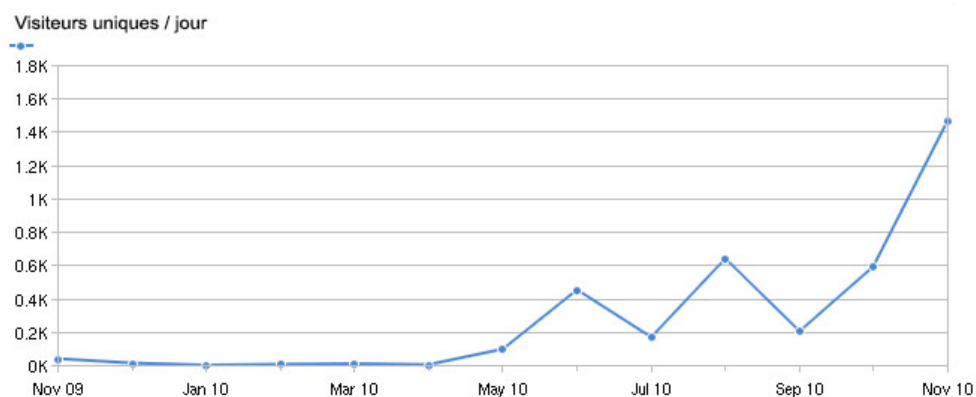
Sa particularité est de produire des *bundles* ou dossiers de fichiers contenant en général, pour un même titre piraté : une version HTML, Word, Mobi, ePub, PDF, ainsi qu'une image de couverture et un fichier texte d'information générale (métadonnées du livre et références de la team⁹).



La qualité globale des fichiers produits par cette team est excellente.

2°) Un nouveau forum et moteur de recherche, qui a su s'imposer en moins d'un an comme l'un des principaux agrégateurs d'ebooks pirates.

La croissance des usages de lecture numérique a pu motiver le développement d'espaces de piratage dédiés aux livres. C'est le cas de ce forum hébergé en Tunisie dont la page Facebook comporte près de 3 000 fans. Selon le site d'évaluation d'audience Sitetrail.com, le forum réalise 160 000 pages vues/mois (pour 1 500 visiteurs uniques/jour), avec un fort taux de croissance depuis 8 mois :



Estimation d'audience d'après Sitetrail.com

Son moteur de recherche fonctionne comme un outil commercial, avec des fiches-livres détaillées (incluant un code à copier-coller dans les forums en cas de partage d'information ou de téléchargement illégal du livre) ainsi qu'un lien d'achat vers Amazon (affiliation).

⁹ On y trouve également, en exergue, la citation suivante de Victor Hugo : « *Le livre, comme livre, appartient à l'auteur, mais comme pensée, il appartient, le mot n'est pas trop vaste, au genre humain. Toutes les intelligences y ont droit. Si l'un des deux droits, le droit de l'écrivain et le droit de l'esprit humain, devait être sacrifié, ce serait, certes, le droit de l'écrivain, car l'intérêt public est notre préoccupation unique, et tous, je le déclare, doivent passer avant nous.* »

Résumés de livres & ebooks - Guide d'Achat

Titre du Livre ou Auteur ou Mot clé:

[Résumés Livres](#) [Recherches Récentes](#) [Derniers Livres Consultés](#) [Top 10 Résumés](#) [Livres au Hasard](#)



1 **Une forme de vie**
Auteur : Amélie Nothomb
Sortie : 18 août 2010



2 **Ni d'Eve ni d'Adam**
Auteur : Amélie Nothomb
Sortie : 29 avril 2009



3 **Hygiène de l'assassin**
Auteur : Amélie Nothomb
Sortie : 6 octobre 2004



Une forme de vie

Résumé / Synopsis :

Détails du Livre :

Auteur : Amélie Nothomb
Sortie : 18 août 2010
Année : 2010
Pages : 168
Type : Broché
Editeur : ALBIN MICHEL
Collection : LITT.GENERALE
Langue : Français

BBCode du Livre (A Insérer dans les forums)

```
[b]Une forme de vie[/b]

[b]Couverture[/b]

[img]
/images/ebooks/2226215174-
une-forme-de-vie.jpg[/img]

[b]Résumé / Synopsis[/b]

[b]Détails du Livre :[/b]

[b]Auteur[/b] : Amélie Nothomb
[b]Sortie[/b] : 18 août 2010
[b]Année[/b] : 2010
[b]Pages[/b] : 168
[b]Type[/b] : Broché
[b]Editeur[/b] : ALBIN MICHEL
[b]Collection[/b] : LITT.GENERALE
```

Autres Livres de l'Auteur : [Amélie Nothomb](#)
Sorties de l'Année : [2010](#)
Ouvrages de L'Editeur : [ALBIN MICHEL](#)
Même Collection : [LITT.GENERALE](#)
Recherches 1. [nothomb](#)
Publicité

Il faut mentionner également le cas particulier de la BeQ (Bibliothèque Electronique du Québec) et des Classiques des sciences sociales¹⁰, qui continue de diffuser, au format PDF et désormais ePub, des classiques de la littérature du XX^e siècle¹¹ dont un certain nombre d'auteurs ne sont pas encore dans le domaine public en France (contrairement au Canada). C'est ainsi le cas de Georges Bernanos, Albert Camus, André Gide, George Orwell ou Antoine de Saint-Exupéry.

Enfin, il est à noter que le SNE est sur le point de rejoindre le dispositif de l'Hadopi, en confiant à un prestataire technique le soin de surveiller sur Internet des listes de titres fournies par les éditeurs. Un processus coûteux et qui ne s'applique pour l'instant qu'aux technologies de peer to peer* et non pas au téléchargement direct (*direct download**).

- ➔ Pour plus d'informations sur la comparaison entre l'offre pirate et l'offre légale, consulter le « Tableau de bord n°1 – Offre numérique » (octobre 2010)
www.lemotif.fr/fr/e-motif/elabz-/l-offre-et-la-demande-de-livres/
- ➔ Pour une étude qualitative sur les uploaders et téléchargeurs, consulter le « Portrait des cyber-pirates du livre » (octobre 2010)
www.lemotif.fr/fr/e-motif/elabz-/portraits-de-pirates/

¹⁰ Site lié à l'université du Québec à Chicoutimi (<http://classiques.uqac.ca/>).

¹¹ <http://beq.ebooksgratuits.com/classiques/index.htm>

Méthodologie de l'étude

Pour conduire notre étude, nous avons identifié les circuits de diffusion des contenus piratés, puis nous avons essayé d'évaluer la quantité de ressources disponibles afin d'en extraire des échantillons analysables. La présente mise à jour suit la même méthode que celle de l'année dernière ; nous la rappelons ici.

Les circuits de téléchargement illégal

Nous ne donnerons pas ici la liste des sites offrant un accès à des fichiers illégaux et consultés dans le cadre de notre étude. Mais plutôt un classement des grandes catégories de sources possibles. Chacune de ces catégories a ses propres logiques de recherche, et donne un accès privilégié à des types de fichiers particuliers.

- 1) ed2k : le P2P* *via* le réseau eDonkey*
- 2) Torrent : le P2P* *via* le réseau torrent*
- 3) DDL : le Direct Download* ou téléchargement direct
- 4) Usenet : les newsgroups ou groupes de discussion
- 5) IRC : l'Internet Relay Chat
- 6) HTTP : les recherches sur le web

Ces différentes sources correspondent à des usages d'Internet très différents.

Utilisateur novice ou occasionnel	HTTP, DDL*
Utilisateur averti	DDL*, eDonkey*, BitTorrent
Utilisateur expérimenté	Usenet, IRC, FTP, <i>trackers*</i> privés

D'une façon générale, et à l'exception de quelques best-sellers très diffusés, il reste assez fastidieux de trouver sur internet la version pirate d'un livre précis. Mais l'amélioration de l'accessibilité est une tendance de fond.

1) Ed2k : le piratage *via* le réseau eDonkey*

eDonkey* est un logiciel de partage de fichiers en peer to peer*, c'est-à-dire de machine à machine. Les sites dits de *dumplinks* ou *indexers* proposent des liens de téléchargement sous la forme d'annuaire thématique : musique, films, séries, jeux vidéo... Des annuaires de plus en plus nombreux comportent une rubrique « ebooks », qui indique les fichiers rendus disponibles gratuitement par les utilisateurs. Nous avons utilisé le logiciel eMule, l'interface eDonkey* la plus utilisée dans le monde. Le logiciel inclut une fonction de recherche qui permet l'utilisation de critères croisés.

Le protocole de recherche utilisé sur eMule est déterminant car il permet de filtrer les résultats avec pertinence. Ainsi une grande partie des fichiers comporte un **mot-clé générique** qui permet de préciser s'il s'agit d'un livre ou d'une BD.

- > pour les livres : ebook[s] french, ebook[s] francais, ebook[s] FR, ebook[s] livre[s], livre[s]
- > pour la bande dessinée : bande dessinee, BD french, BD francais, comics French, BDFR

L'utilisation de cette méthode introduit nécessairement un biais du fait que l'internaute n'effectue que rarement une recherche sur un mot-clé générique, mais plutôt sur le titre d'un livre ou un nom d'auteur. Toutefois, elle reste pertinente car l'on constate que la majorité des fichiers sont « taggués » (identifiés par un tel mot-clé) de cette manière, ne serait-ce qu'au moins une des versions pirates d'un même livre, lorsqu'il en existe plusieurs. Afin d'obtenir un échantillon de taille raisonnable à analyser, nous avons effectué ces recherches sur les principaux serveurs ed2k*, en ne tenant compte que des fichiers réellement téléchargeables (on observe ainsi qu'en deçà de 2 sources accessibles sur plusieurs semaines, le fichier est difficilement accessible.)

A ces mots-clés ont été associés les extensions de formats de fichiers les plus répandus : .pdf (format PDF), .epub (ePub), .doc (format Word), .txt (format texte brut), .djvu, et les fichiers compressés .rar (WinRar), .zip (WinZip) et .cbr (Comic Book Rar).

Nous avons par ailleurs contrôlé ces recherches avec, selon le même protocole, des recherches sur l'agrégateur Figator.

Werber, Bernard - Le Livre Secret Des Fourmis

Cache

Extension: pdf

Availability: 5

Complete Sources: 1

Size: 13.6 Mb

Location: unknown

This file has been seen in peer to peer networks with the following names:

Werber, Bernard - Le Livre Secret Des Fourmis.pdf

Capture d'écran d'un exemple de fiche d'ebook pirate référencé sur Figator

2) BitTorrent : le P2P* via le réseau torrent*

Torrent* est également un protocole d'échanges de fichiers en peer to peer*. Le fonctionnement est toutefois différent car il suppose, pour un pirate qui souhaite diffuser une *release**, de créer un fichier torrent* (*tracker**) et de le faire référencer sur un annuaire spécialisé. Cette création et ce référencement, s'ils sont simples, rendent tout de même la procédure plus compliquée que *via* eDonkey.

Pour effectuer une recherche sur les torrent*, il faut donc accéder à des annuaires ou agrégateurs qui répertorient les liens et les classent en différentes catégories thématiques, dont les ebooks, mais dont l'ergonomie de recherche est rarement performante.

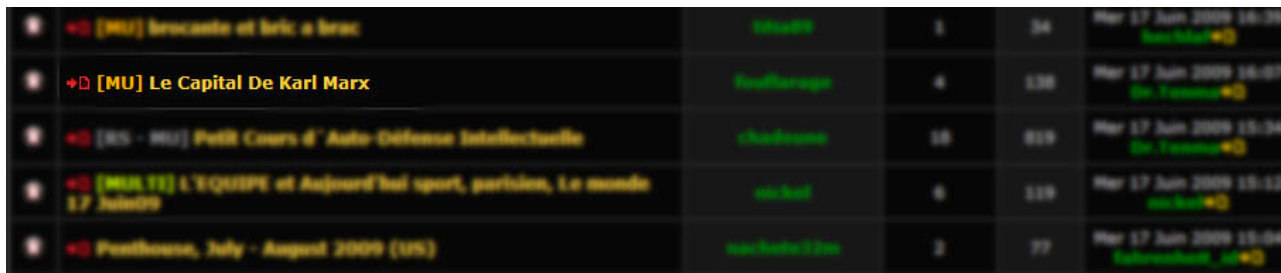
La méthode d'identification des fichiers utilisée est donc l'extraction exhaustive des fichiers inclus dans ces catégories « ebooks » des principaux agrégateurs.

3) DDL : le *direct download* ou téléchargement direct

Le *direct download**¹² est en progression exponentielle depuis quelques années, car il offre à la fois une alternative sécurisée face à la surveillance du piratage, et des vitesses de téléchargement souvent plus élevées.

L'apparition de sites de stockage comme MegaUpload, RapidShare ou le service d'hébergement de Free, qui proposent des espaces de plusieurs gigas pour y déposer des fichiers dont la légalité n'est que rarement vérifiée, a contribué au développement de ce mode de téléchargement. Des sites spécialisés, en général sous la forme de forums nécessitant une inscription, fournissent des listes de liens de téléchargement direct renouvelés quotidiennement. Ces liens sont généralement standardisés : le titre commence par le nom de l'hébergeur entre crochets, suivi par le titre du livre proposé.

Exemple : [MU] *Le capital de Karl Marx* : il s'agit d'un lien de téléchargement du livre *Le Capital*, hébergé sur MegaUpload. [MULTI] signifie que les liens de téléchargement sont disponibles sur plusieurs sites de stockage différents.



Capture d'écran d'un forum de warez

La méthode d'identification des fichiers est la même que pour BitTorrent : l'extraction des fichiers contenus dans les rubriques « ebooks » de principaux sites et forums de référencement.

Des moteurs de recherche spécifiques aux hébergeurs de DDL* restent actifs, comme www.filecrop.com ou encore www.filestube.com. Toutefois, ces moteurs ne retournent que les résultats contenant le mot-clé de la recherche dans le titre du fichier, ce qui est de fait très limitatif. Par ailleurs, les fichiers sont filtrés très régulièrement et les contenus illégaux supprimés, ce qui augmente considérablement le taux de liens de téléchargement référencés mais non accessibles.

Il existe un autre moyen plus surprenant de trouver des fichiers sur le web. Il consiste à utiliser les fonctionnalités de recherche avancée de Google. On appelle parfois ces dernières « Google hacks » car elles peuvent, dans certains cas, permettre l'accès à des données censées être non publiques. Par exemple, la ligne suivante indexe des listes de sites proposant des livres en téléchargement :

```
intitle:"index.of" ebooks, intitle:"index of" +("/ebooks"|"book") +(chm|pdf|zip)
```

Cette dernière méthode reste anecdotique, car utilisée seulement par les plus expérimentés, et nous n'en avons pas tenu compte dans nos recherches de fichiers.

4) Usenet : les newsgroups ou groupes de discussion

Plus connus en France sous le nom de « newsgroups », il s'agit d'un réseau de forums de discussion, existant depuis 1979, et dont Google Groups n'est que la partie grand public la plus visible aujourd'hui. A l'origine développés pour diffuser du texte brut, les newsgroups sont réputés pour leur capacité à véhiculer des fichiers binaires*, par exemple des fichiers de musique, des films ou des ebooks.

Les noms des groupes de téléchargement d'ebooks francophones contiennent généralement le mot « binaries » (fichiers binaires*) ainsi que le mot « ebook » et « french ». Des moteurs de recherche spécifiques, sur le web, comme www.binsearch.info, permettent de trouver plus facilement les fichiers diffusés sur Usenet.

La méthode d'extraction a été identique à celle du réseau ed2k, c'est-à-dire l'utilisation de ces mots-clés génériques.

¹² Le DLL est en réalité une forme de téléchargements sur le web puisqu'il s'agit de liens directement accessibles depuis un navigateur web. Toutefois ses spécificités justifient une catégorie indépendante.

5) IRC : l'Internet Relay Chat

L'Internet Relay Chat ou IRC est un protocole de communication par texte instantané.

Le téléchargement par IRC est réservé à une minorité d'utilisateurs chevronnés, car il requiert des compétences informatiques dépassant largement la moyenne de celles de la population. Il suppose la connexion à un canal (« channel »), nécessitant souvent une inscription préalable¹³. Il s'agit ensuite d'opérer un certain nombre de commandes pour la recherche et le téléchargement. Devant l'impossibilité de trouver des « annuaires » d'ebooks illégaux sur IRC, la méthode utilisée a été la même que pour ed2k et Usenet. Plus précisément, on se reportera aux commandes de recherche comme @find ou @search suivies du titre du livre et aux commandes de téléchargement sous la forme : !serveur ebook - Bernard Werber - Les Fourmis [Html].rar ::INFO:: 167.0KB)

6) HTTP : les consultations en visionnage ou « streaming » sur le web

Mis à part les recherches spécifiques sur un moteur de recherche comme Google, il est possible de trouver des ebooks pirates disponibles en lecture en ligne (avec éventuellement un lien de téléchargement) sur des sites de publication spécialisés, pour lesquels nous avons utilisé la même méthode que pour ed2k.

Ayant fait largement parler de lui auprès du grand public quand Barack Obama l'a utilisé pour diffuser des documents pour sa campagne électorale, le site **Scribd** (www.scrib.com), créé en 2007 aux États-Unis, permet de publier ses textes dans une base de données accessible à tous. Les mêmes remarques peuvent être faites sur **WattPad** (www.wattpad.com), un logiciel de lecture d'ebooks sur téléphones portable lancé en 2006, lui aussi a été très critiqué pour héberger sans accord préalable des fichiers sous droits.

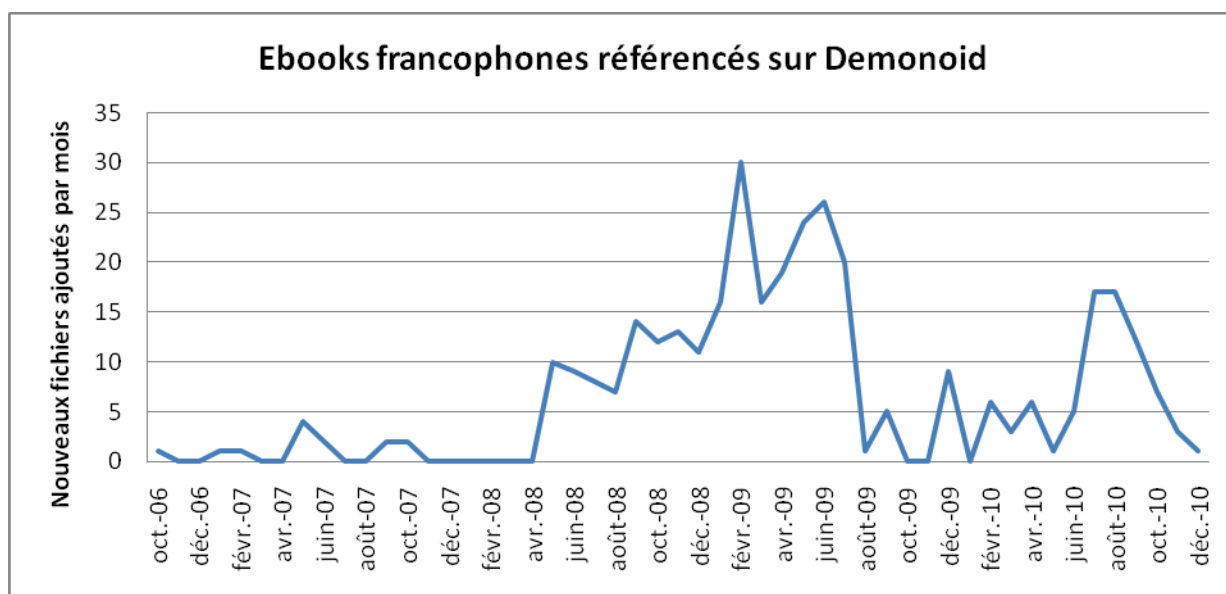
Si tous ces sites suppriment régulièrement des documents illicites, à la demande des ayants droit, il reste encore une poignée d'ouvrages francophones disponibles au visionnage en toute illégalité. Toutefois les fichiers disponibles directement par HTTP ne constituent qu'une part très marginale de l'ensemble : un internaute novice qui effectue une recherche d'ebook pirate sur un moteur de recherche obtiendra d'abord comme résultats les liens vers du téléchargement direct ou du peer to peer.

¹³ Les *channels* anglophones les plus populaires portent les noms évocateurs de #ebooks, #bookz et #audiobooks.

Focus — Déclin du peer to peer au profit du téléchargement direct

Les études sur les échanges illégaux constatent que le peer to peer, *via* ed2k comme le réseau torrent, est en déclin au profit du téléchargement direct. Ainsi, l'une d'entre elles¹⁴ montre que le trafic des deux principaux sites de *direct download** (RapidShare et MegaUpload) dépasse celui de Facebook et représente plus de 2 % du trafic mondial sur Internet. Une autre analyse révèle qu'au début de l'année 2010 on constatait (tous produits culturels confondus) 28 téléchargements par direct download* pour 100 téléchargements par torrent ; en juillet 2010, le rapport passe à 78 téléchargements par DDL pour 100 torrent¹⁵.

Fichiers référencés sur le moteur de recherche de torrent Demonoid

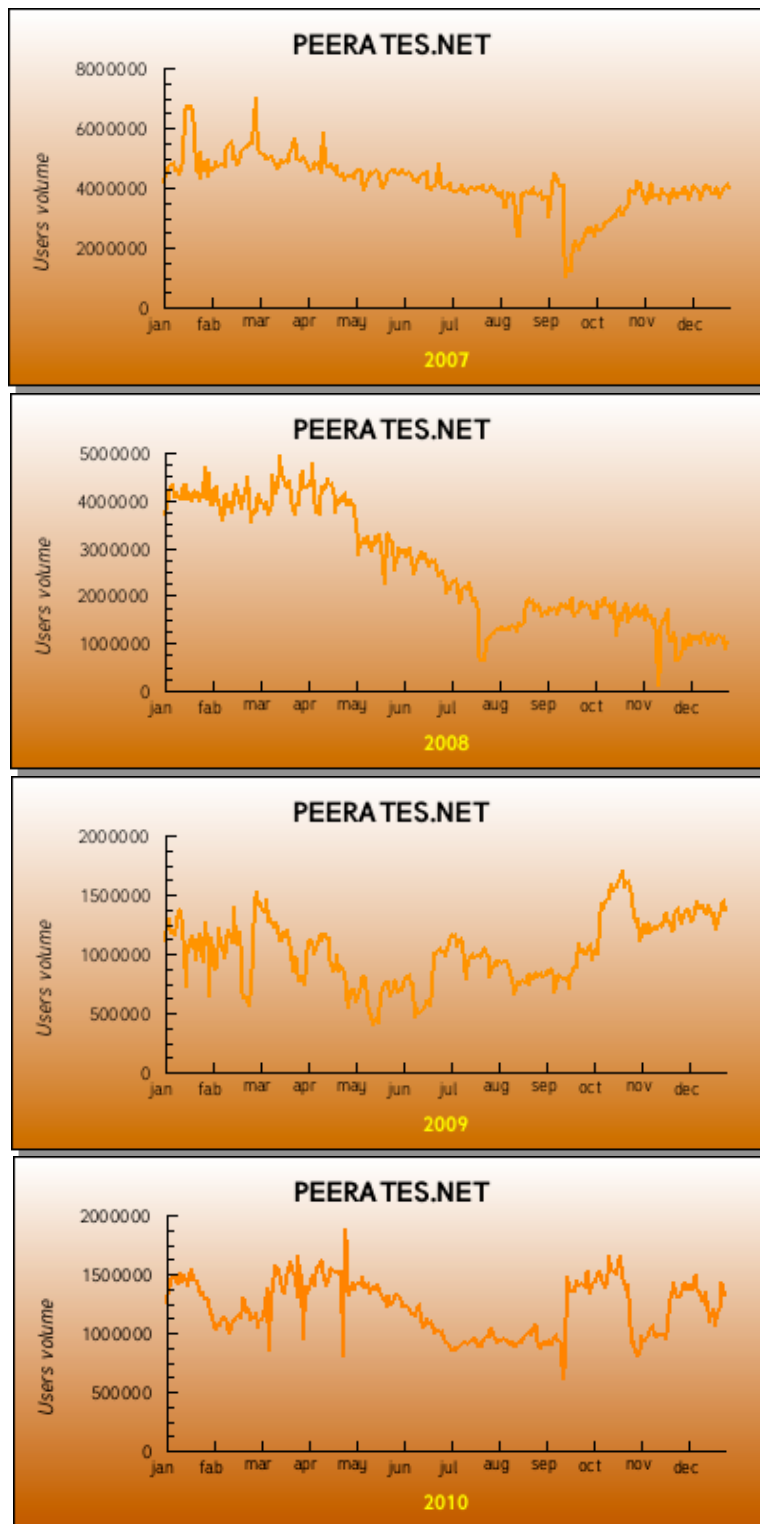


Source : Demonoid.com, janvier 2011.

Ces constatations sont particulièrement vérifiées pour le livre, pour lequel la relative petite taille des fichiers (comparée au poids d'un album de musique, d'un jeu vidéo ou d'un film) rend ce mode de téléchargement plus pertinent. Les diagrammes ci-dessous, fournis par le site d'analyse des échanges ed2k Peerates, confirment cette tendance, puisqu'on est passé d'un volume d'environ 4,5 millions d'utilisateurs en 2007 à 1 million en 2010 :

¹⁴ <http://www.sandvine.com/>

¹⁵ <http://www.plugngeek.net/le-telechargement-direct-explose>



Evolution annuelle du nombre de clients eDonkey connectés chaque jour, calculée sur la base de deux relevés par jour.

Une évaluation globale de l'offre

Dans « EbookZ 1 », nous avons tenté une estimation du nombre de fichiers illégaux disponibles.

Après trois nouveaux mois de recherches sur un très grand nombre de fichiers illégaux à disposition des internautes dans les différents circuits de diffusion, nous estimons que sont disponibles :

- ❖ **2 000 à 3 000 titres** de ebooks (*i.e.* livres, par différence avec les BD) ;
- ❖ 30 000 à 35 000 titres de bandes dessinées, dont environ **6 000 à 7 000 titres** sont réellement accessibles (la majorité des fichiers sont partagés en peer to peer avec très peu de sources de téléchargement).

Soit un total plutôt conservateur d'environ 8 000 à 10 000 titres en cumulant livres et BD (sont exclus les ouvrages disponibles seulement sur des réseaux privés, accessibles par une connexion sécurisée, et dont il est impossible d'évaluer la quantité), ce qui est très faible au regard du nombre de titres papier disponibles.

	Titres piratés	Titres disponibles en offre légale papier	% des titres piratés
Livres + BD	8 000-10 000	619 800	1,2 %-1,6 %

En ce qui concerne le peer to peer*, la majorité des fichiers disponibles ne sont que difficilement accessibles, car peu partagés par les internautes (entre 0 et 1 source). Le téléchargement illégal de livres est un piratage résiduel dans un secteur qui n'a pas encore connu de commercialisation numérique massive. Ce taux de piratage progresse néanmoins avec la mise à disposition d'ebooks au format numérique, nécessitant une mise à jour régulière de cette étude.

Comparaison avec 2009

Nous avons dénombré, dans « EbookZ 1 », 1 000 à 1 500 titres d'ebooks et 3 000 à 4 500 titres de BD (ainsi que 200 à 300 titres de livres audio), soit 4 000 à 6 000 titres en tout. On voit que l'offre en 2010 a augmenté de presque 70 % dans sa fourchette haute.

En valeur absolue, le volume global reste faible au regard de l'offre légale papier, mais à l'heure d'aujourd'hui il nous faut raisonner également en tenant compte de l'offre légale numérique, évaluée entre 60 000 et 70 000 titres ; il resterait cependant à savoir quels sont les titres de cette offre légale numérique qui se trouvent piratés.

Echantillonnage

A partir de ces observations, nous avons pu établir des ensembles ou échantillons¹⁶ que nous avons analysés en détail afin de comprendre la nature exacte de l'offre illégale. Nous avons retenu le nombre de fichiers suivants :

	Nombre de fichiers
Ebooks	769
BD	246
Total	1 015

Il est à noter que, pour cette mise à jour, nous avons utilisé une méthode chronologique en ce qui concerne le téléchargement direct (seul circuit de diffusion sur lequel les fichiers peuvent être datés facilement), ce qui signifie que nous avons traité en priorité les fichiers mis à disposition le plus récemment (nous sommes remonté jusqu'en septembre 2010).

Indépendamment des nouveautés, cette méthode reste pertinente pour juger de l'ensemble des fichiers illégaux, d'une part parce que ces derniers ont une « date de péremption » liée à leur suppression par les hébergeurs de *direct download** lorsque les liens ne sont plus actifs, ce qui oblige les pirates à les re-uploader régulièrement ; et d'autre part parce qu'elle ne concerne pas les fichiers partagés en peer to peer.

Comme pour « EbookZ 1 », ces échantillons ont été récupérés selon une logique d'entonnoir : nous avons commencé par relever les fichiers disponibles facilement (cf. la classification des types de téléchargeurs page 10), c'est-à-dire sur le réseau eDonkey et en BitTorrent ainsi que sur les principaux annuaires de *direct download** et marginalement sur les sites de publication en HTTP. Puis nous avons affiné en utilisant les sources plus alternatives comme les newsgroups et l'IRC.

Nous nous sommes donc mis à la place d'un internaute moyennement expérimenté, afin de voir ce qu'il était possible de trouver sans prérequis informatiques importants.

La taille des ensembles a été déterminée de façon à conserver le meilleur rapport possible entre une nécessaire représentativité et la complexité d'analyser un grand nombre de fichiers. Au total, les quantités de livres piratés sont suffisamment faibles pour que les ensembles considérés soient assez représentatifs : ainsi, pour les livres, l'échantillon représente entre 25 % et 38 % de l'estimation du volume global d'ouvrages piratés disponibles sur Internet.

Les échantillons ont toujours tendance à sous-représenter certaines catégories de livres, trouvables en grandes quantités sur des sites spécialisés, voire privés, mais qui n'apparaissent que dans une moindre mesure lors de recherches généralistes. C'est surtout le cas des publications professionnelles (ouvrages à destination de l'entreprise, par exemple d'information financière, marketing ou scientifique), d'informatique et de jeux de rôles.

Au final, la répartition par circuits des corpus de fichiers que nous avons étudiés est la suivante :

	Emule	Torrent	DDL	Usenet	Autre
Livres	15,6 %	6,5 %	74,3 %	2 %	1,4 %
BD	24,8 %	2,4 %	66,2 %	5,3 %	1,2 %
Total	17,8 %	5,6 %	72,4 %	2,8 %	1,4 %

¹⁶ Les fichiers ont été recueillis pendant une période de trois mois allant du 1^{er} octobre 2010 au 1^{er} janvier 2011.

La grande différence avec « EbookZ 1 » tient donc à l'importance prise par le téléchargement direct (DDL), au détriment de l'échange de fichiers en peer to peer. On se rend compte qu'il y a moins de fichiers disponibles facilement et immédiatement sur les réseaux peer to peer ; celui-ci représente moins d'un quart des fichiers d'ebooks pirates. Les autres modes de téléchargement, sans surprise, restent marginaux. La répartition que nous avons constatée dans le tableau ci-dessus est cohérente avec les études sur la répartition des circuits de téléchargement.

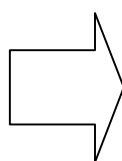
Il est à noter que les chiffres ci-dessus ne reflètent pas la diffusion de l'ensemble des textes piratés, mais de ceux qui sont accessibles le plus facilement, selon la méthode utilisée pour créer les échantillons de l'étude. Ils ont donc tendance à sous-valoriser certains circuits moins accessibles aux profanes, à l'instar des newsgroups (essentiellement pour les bandes dessinées). Par ailleurs, ils ne tiennent compte que du téléchargement et pas du visionnage en ligne ou *streaming*.

Catégories éditoriales

Afin d'affiner les comparaisons avec le marché du livre papier, et de permettre aux professionnels de disposer de références connues, nous avons choisi, pour cette mise à jour, de suivre la typologie établie par le SNE¹⁷.

Nous avons donc supprimé la distinction opérée entre « Genres » et « Thèmes » dans « EbookZ 1 », et avons retenu 22 catégories et sous-catégories (parmi les 10 catégories principales de la nomenclature en usage dans la profession¹⁸.)

1-ENSEIGNEMENT
2-SCIENCES TECHNIQUES ET MEDECINE
3-INFORMATIQUE
4-SCIENCES HUMAINES ET SOCIALES
5-PHILOSOPHIE
6-RELIGION ET ESOTERISME
7-DICTIONNAIRES ET ENCYCLOPEDIES
8-ROMANS
9-ROMANS POLICIERS
10-ROMANS EROTIQUES
11-ROMANS SF ET FANTASTIQUES
12-THEATRE ET POESIE
13-DOCUMENTS ET ESSAIS
14-HUMOUR
15-MEMOIRES ET BIOGRAPHIES
16-JEUNESSE
17-BD ALBUMS
18-BD COMICS
19-BD MANGA
20-BEAUX ARTS
21-PRATIQUE ET LOISIRS
22-TOURISME



1-ENSEIGNEMENT
2-SCIENCES ET TECHNIQUES, MEDECINE, GESTION
3-SCIENCES HUMAINES ET SOCIALES
4-RELIGION ET ESOTERISME
5-DICTIONNAIRES ET ENCYCLOPEDIES
6-LITTERATURE
7-DOCUMENTS, ACTUALITE, ESSAIS
8-JEUNESSE
9-BANDE DESSINEE
10-BEAUX LIVRES ET LIVRES PRATIQUES

¹⁷ Pour connaître le détail de ces catégories, nous renvoyons le lecteur aux repères statistiques du SNE : <http://www.sne.fr/pages/infos-et-documentation/publications.html>

¹⁸ Il y en a 12 mais nous en excluons Cartes & Atlas et Ouvrages de documentation.

Résultats

Il est nécessaire de rappeler que les résultats de notre étude doivent être pris avec précaution, compte tenu de la volatilité des échanges sur Internet et de la difficulté d'établir avec certitude des échantillons représentatifs sur des données illégales.

Analyse des fichiers

Qualité

Pour juger de la qualité d'un fichier, une dizaine de caractéristiques ont donné lieu à une évaluation globale. Les caractères du texte sont-ils flous ? Le texte est-il imprimable ? Respecte-t-il la mise en page originale ? Les photos ou illustrations ont-elles beaucoup perdu à la compression ? Y a-t-il des pages ou des fragments de page manquants ? Les pages sont-elles alignées correctement ? Le fichier est-il bien compressé ? etc.

L'évaluation correspond à un chiffre de **1 à 5** :

- 1 : fichier le plus bas de gamme, en général un simple texte au format brut, sans mise en page, difficilement lisible ou incomplet.
- 2 : fichier au format traitement de texte (Word ou Open Office), très légèrement remis en page, ou PDF de très mauvaise qualité.
- 3 : fichier moyen, en général un PDF au format image dont la résolution ne permet pas l'impression.
- 4 : fichier de très bonne qualité, auquel il manque seulement un ou deux critères de confort de lecture (par exemple : recherche en mode texte dans le cas d'un essai).
- 5 : fichier au format PDF (image ou en texte libre) ou ePub d'excellente qualité, se rapprochant au maximum de l'original et offrant un confort de lecture optimal¹⁹.

Remarque : contrairement à la musique et à la vidéo, le taux de *fakes** pour les ebooks est toujours extrêmement bas, moins de 1 %.

	1	2	3	4	5	Note moyenne
Livres	2,6 %	8,7 %	21,5 %	22,1 %	45,2 %	3,9
BD	0 %	2,2 %	6,5 %	52,7 %	38,7 %	4,3
Total	1,2 %	7,2 %	18 %	29,1 %	43,7 %	4

Notre étude montre que la **qualité globale moyenne des fichiers reste à 4 sur 5**. 73 % des fichiers (67,3 % des livres et 91,4 % des BD) ont une qualité supérieure ou égale à 4. Ce qui va à l'encontre d'une idée reçue selon laquelle les fichiers illégaux sont de mauvaise qualité.

Par rapport à « EbookZ 1 », on observe une légère amélioration de la note moyenne (3,9 pour les ebooks contre 3,8), qui s'explique par le développement de l'offre légale et la multiplication du nombre de fichiers crackés.

¹⁹ On peut se demander dans quelle mesure il s'agit de fichiers provenant directement de versions numériques « crackées » ; on a pu le constater dans le cas d'ouvrages de STM publiés chez des éditeurs comme Eyrolles.

Explication des formats :

- **PDF image/texte** : fichiers au format PDF d'Adobe, avec ou sans la possibilité de lire le contenu en mode texte (donc de copier/coller des passages). L'essentiel de ces fichiers provient de scans réalisés manuellement depuis des ouvrages papier. Une faible quantité (que l'on peut évaluer cette année encore à moins de 4 %²⁰) provient du piratage de fichiers numériques préexistants disponibles en offre légale, et dont il est quasi impossible de définir à quel niveau s'est effectuée la fuite (acheteur, service de presse de l'éditeur, correcteur, auteur, etc.)
- **Word** : fichiers au format .doc ou .rtf (dans quelques très rares exceptions, au format .odt d'Open Office).
- **TXT** : fichiers au format de texte brut, sans mise en forme.
- **DjVu** : format de fichier destiné à l'archivage de documents numériques, créé par AT&T aux États-Unis²¹. Les fichiers sont lus par un logiciel très léger (moins de 3 Mo) et ergonomique. Il permet une compression optimisée en termes de taille et de qualité.
- **IMG** : fichiers dans un format image, généralement en JPEG. Dans le cas des BD, ce format signifie donc un fichier découpé par planches.
- **MULTI** : dossier contenant plusieurs fichiers d'un même titre aux formats différents (en général, une version epub + pdf + doc).

	PDF Image	PDF Texte	ePub	Word	TXT	DjVu	IMG	MULTI
Livres	36,5 %	42,5 %	3%	2,2 %	0,2 %	1,1 %	0,9 %	13,5 %
BD	39,8 %	0 %	0 %	0 %	0 %	0 %	60,2 %	0 %
Total	37,5 %	29,2 %	2,1 %	1,5 %	0,1 %	0,8 %	19,5 %	9,3 %

La grande majorité des fichiers d'ebooks (79 %) restent au **format PDF**, le plus pratique pour une lecture multi-supports. Les formats textes, tous confondus²², représentent **61,4 % des livres** piratés.

La nouveauté entre la fin de 2009 et la fin de 2010 est **l'apparition de fichiers ePub pirates**, qui étaient quasi-inexistants avant l'été 2009. En considérant que presque chaque dossier MULTI contient une version ePub, on peut donc considérer que ce format représente aujourd'hui plus de **16 % (3 % + 13,5 %) des ebooks pirates**. Ce taux continuera certainement d'augmenter avec le développement de l'offre légale et des supports de lecture adaptés.

En ce qui concerne la BD, la répartition entre PDF et IMG (images) reste sensiblement la même qu'en 2009. A noter, comme dit plus haut, qu'il n'est pas tenu compte ici du visionnage de BD en streaming, par exemple sur des sites de scantrad*.

²⁰ Dans notre échantillon, on peut repérer ce type d'ouvrages soit, très rarement, par la présence de *watermarks* (tatouages numériques) sur le fichier, soit par le rendu clairement professionnel de ce dernier.

²¹ <http://fr.wikipedia.org/wiki/DjVu>

²² PDF texte + Word + TXT + ePub + MULTI. DjVu étant un cas à part. L'usage de ce dernier a fortement baissé par rapport à notre première étude, ce qui s'explique notamment par la disparition de l'un des principaux agrégateurs d'ebooks pirates en direct download*, qui comportait beaucoup de fichiers djvu.

Taille

La taille des fichiers d'ebooks est beaucoup moins standardisée que celle des fichiers de musique (qui eux sont souvent en format mp3 et compris entre 3 et 5 Mo), de films et de jeux (conçus pour tenir sur un CD de 700 Mo).

	< 1 Mo	1 à 4 Mo	5 à 9 Mo	10 à 49 Mo	50 à 249 Mo	> 250 Mo	Moyenne
Livres	13,8 %	33,7 %	11,7 %	26,1 %	14,6 %	0,2 %	20,1 Mo
BD	0,0 %	0,0 %	4,3 %	66,7 %	28,2 %	0,9 %	39,4 Mo
Total	11,1 %	27,2 %	10,3 %	33,9 %	17,2 %	0,3 %	23,8 Mo

Les livres sont en général des fichiers de petite taille comparée à un album de musique ou une vidéo. Toutefois, contrairement à beaucoup d'idées préconçues, l'étude confirme que la taille moyenne est assez élevée : **20,1 Mo** pour les livres, **39,4 Mo** pour les bandes dessinées. Cela traduit, pour les ebooks, le grand nombre d'ouvrages illustrés (surtout scolaires et pratiques) scannés au format image qui pèsent beaucoup plus lourd qu'un format texte²³.

Cependant, on constate que la taille moyenne des fichiers d'ebooks a baissé, puisqu'elle était à 29,7 Mo dans notre première étude. Cela s'explique par le développement des fichiers textes, notamment au format epub, très léger (sur notre échantillon, la moyenne des fichiers epub est de 500 Ko).

La forte différence observée pour la BD (de 79,3 Mo à 39,4 Mo) avec « EbookZ 1 » est d'ordre méthodologique : nous n'avons pas tenu compte du poids des *bundles* ou dossier contenant plusieurs albums d'une même série. La taille n'est donc ici que celle des albums individuels.

Détail des tailles moyennes de fichiers par catégorie éditoriale :

1-ENSEIGNEMENT	16,2 Mo
2-SCIENCES TECHNIQUES ET MEDECINE	29,6 Mo
3-INFORMATIQUE	20,4 Mo
4-SCIENCES HUMAINES ET SOCIALES	18,5 Mo
5-PHILOSOPHIE	9,7 Mo
6-RELIGION ET ESOTERISME	8,7 Mo
7-DICTIONNAIRES ET ENCYCLOPEDIES	59 Mo
8-ROMANS	2,1 Mo
9-ROMANS POLICIERS	1,8 Mo
10-ROMANS EROTIQUES	0,5 Mo
11-ROMANS SF ET FANTASTIQUES	4,5 Mo
12-THEATRE ET POESIE	5,4 Mo
13-DOCUMENTS ET ESSAIS	17,6 Mo
14-HUMOUR	4 Mo
15-MEMOIRES ET BIOGRAPHIES	8,3 Mo
16-JEUNESSE	25 Mo
17-BD ALBUMS	35,7 Mo
18-BD COMICS	77,9 Mo
19-BD MANGA	55,4 Mo
20-BEAUX ARTS	78,5 Mo
21-PRATIQUE ET LOISIRS	42,6 Mo

²³ En considérant seulement les formats textes (PDF texte, Word, TXT et HTML), la moyenne des fichiers est **5 Mo**.

Répartition des releasers*

	Rang 1	Rang 2	Rang 3	Rang 4	Rang 5
% des fichiers	6 %	4,4 %	1 %	0,6 %	0,4 %
% des fichiers avec releaser identifié	44,2 %	32,6 %	8 %	4,3 %	2,9 %

Ce tableau se lit ainsi : 6 % de l'ensemble des fichiers disponibles proviennent du *releaser** le plus productif (de rang 1), 4,4 % du deuxième, etc. Cela permet d'évaluer la part des équipes et des gros utilisateurs dans le piratage des livres : livres et BD confondus, **2 releasers* concentrent ainsi plus de 10 % des créations de fichiers illégaux.**

A noter qu'il n'est souvent pas explicitement possible de connaître le *releaser**, à défaut de la présence de l'un des 3 éléments suivants : identification de la team dans le nom du fichier, dans le post du forum comportant le lien à télécharger, ou dans un fichier texte spécial livré avec l'ebook pirate (souvent un .txt comportant les informations sur le *releaser**). La **majorité des fichiers (86,4 %) n'ont pas de releaser* clairement identifié.**

La concentration des fichiers chez les deux premiers *releasers** (76,8 %) montre l'existence de deux équipes de pirates, très organisées, qui produisent beaucoup de fichiers repris par la plupart des agrégateurs.

La qualité moyenne des fichiers de ces deux *releasers** est de 4,8 sur 5, ce qui prouve leur objectif d'abord qualitatif (même si ces teams ont aussi un objectif de production soutenue et régulière). La totalité des fichiers de ces équipes est au format PDF texte et/ou ePub.

Analyse des ouvrages

Top 20 des auteurs ayant le plus de titres piratés (livres)

Ce tableau n'indique pas quels sont les auteurs les plus téléchargés illégalement dans l'absolu, mais quels sont ceux qui ont le plus de titres piratés, d'après nos échantillons.

Titres piratés			Titres différents piratés		
	Auteur	Titres piratés		Auteur	Titres différents piratés
1	Bernard Werber	22	1	Robin Hobb	21
2	Robin Hobb	21	2	Bernard Werber	15
3	Isaac Asimov	16	3	Isaac Asimov	13
4	Frédéric Lenormand	12	4	Frédéric Lenormand	12
5	Fred Vargas	11	5	Georges Simenon	8
6	Amélie Nothomb	10	6	Amélie Nothomb	7
7	J. K. Rowling	9	7	Frédéric Beigbeder	7
8	Douglas Adams	9	8	Fred Vargas	6
9	Frédéric Beigbeder	8	9	Douglas Adams	6
10	Maxime Chattam	8	10	J. K. Rowling	6
11	Georges Simenon	8	11	Robert Heinlein	5
12	Agatha Christie	7	12	Sherrilyn Kenyon	6
13	Dan Brown	6	13	Bryan Perro	4
14	J. R. R. Tolkien	6	14	Raymond E. Feist	4
15	Sherrilyn Kenyon	6	15	Magali Ségura	4
16	Robert Heinlein	5	16	Ian Fleming	4
17	Stephenie Meyer	5	17	Stephenie Meyer	4
18	David Bry	5	18	Stephen King	4
19	Raymond E. Feist	4	19	Agatha Christie	4
20	Michel Houellebecq	4	20	Gilles Deleuze	4

Quelques observations :

- ❖ **Continuité des best-sellers** : plus de la moitié des auteurs du classement font toujours partie des auteurs de best-sellers en France²⁴. Leur présence traduit logiquement leur popularité auprès du grand public.
- ❖ **Forte représentation de certaines catégories d'auteurs** :
 - Les auteurs de **science-fiction et fantastique**²⁵ représentaient 25 % des auteurs du classement dans « EbookZ 1 ». Ils en représentent ici **60 %** (en tenant compte des titres différents ou pas).
 - Les auteurs de **romans policiers**²⁶ représentent **25 %** des auteurs du classement (20 % en titres différents).

²⁴ Dont : Bernard Werber, Fred Vargas, Amélie Nothomb, J. K. Rowling, Frédéric Beigbeder, Maxime Chattam, Agatha Christie, Dan Brown, Stephenie Meyer, Michel Houellebecq.

²⁵ Bernard Werber, Robin Hobb, J.K. Rowling, Isaac Asimov, Stephenie Meyer, Ray Bradbury, Raymond E. Feist, Robert Heinlein, Sherrilyn Kenyon, Douglas Adams, J.R.R. Tolkien, David Bry, Magali Ségura, Bryan Perro, Stephen King.

²⁶ Fred Vargas, Maxime Chattam, Georges Simenon, Agatha Christie, Dan Brown.

- Les auteurs philosophes, qui représentaient 25 % du classement dans « EbookZ 1 », ont quasiment disparu, à l'exception de Gilles Deleuze qui passe de la 1^{ère} à la 20^e place.

On retrouve 30 % (35 % en titres différents) des auteurs présents dans le classement de « EbookZ 1 ».

Top 20 des titres les plus piratés (livres)

Le classement des titres les plus référencés²⁷ recoupe partiellement la liste ci-dessus.

1	Apocalypse Bébé	Virginie Despentes
2	Tout couscous	Sophie Brissaud
3	Dracula l'immortel	Dacre Stoker
4	Bilbo le Hobbit	J. R. R. Tolkien*
5	La Carte et le Territoire	Michel Houellebecq*
6	Da Vinci Code	Dan Brown*
7	Les Meilleurs Farcis natures	Stéphanie Blanquet
8	Le Meilleur des Mondes	Aldous Huxley
9	Le Voleur d'ombres	Marc Levy
10	Encyclopédie de la cuisine au micro-ondes	Laura Landra
11	Peut-on jouer au frisbee avec une tong ?	Christian Camara
12	Le Président des riches	Monique et Michel Pinçon-Charlot
13	Debout les morts	Fred Vargas*
14	Autre-monde 3 - Le cœur de la terre	Maxime Chattam*
15	Le Sexe pour les nuls	Ruth K. Westheimer
16	L'Ultime secret	Bernard Werber*
17	L'Arbre des possibles et autres histoires	Bernard Werber*
18	Twilight tome 4 – Révélation	Stephenie Meyer*
19	Le Petit Prince	Antoine de Saint-Exupéry
20	Harry Potter et les reliques de la mort	J. K. Rowling*

* Fait partie du classement des auteurs ayant le plus de titres piratés

Quelques constatations :

- **Prédominance des ouvrages très grand public** et des best-sellers, comme dans « EbookZ 1 ».
- **Prime à la nouveauté** : au sein des ouvrages grand public et/ou best-sellers, 4 sont des nouveautés de moins de 6 mois à la date de l'enquête : *Apocalypse Bébé*, *La Carte et le Territoire*, *Le Voleur d'Ombres* et *Le Président des riches*. On remarquera que *Apocalypse Bébé* et *Le Voleur d'ombres* disposaient d'une offre numérique légale au moment de la collecte de nos données. *La Carte et le Territoire*, cas un peu à part, avait été piraté avec le « buzz » médiatique qu'on connaît. Quant au *Président des riches*, il était disponible gratuitement (en lecture en ligne, au format HTML) sur le site de son éditeur, Zones. La non-disponibilité en offre légale numérique de ces livres ne semble pas expliquer dans ce cas-là leur piratage.
- 4 titres sont des « **valeurs sûres** » du piratage et étaient déjà présents dans le top 20 de 2009 : *Harry Potter*, *Le Sexe pour les nuls*, *Le Petit Prince* et *Twilight*. Aux auteurs de ces livres on peut ajouter J.R.R. Tolkien et Bernard Werber qui eux aussi figuraient l'année dernière.
- L'examen de la convergence des tops 50 (pour l'année 2010) de *Livres-Hebdo* et de l'offre disponible au téléchargement illégal de nos modèles montre un taux de piratage de 6 %²⁸. Ce

²⁷ Il s'agit des titres référencés pendant les 3 mois d'observation sur les circuits de diffusion illégaux que nous avons explorés, classés par nombre de liens de téléchargements disponibles.

pourcentage peut paraître faible au regard de l'importance croissante prise par les nouveautés dans le piratage, mais il faut noter que le classement *Livres-Hebdo* utilisé ne recouvre que le top 50 des parutions de l'année passée, laissant donc de côté les long-sellers ainsi que les nouveautés les plus récentes. Par ailleurs, nous n'avons pas inclus la BD dans cette analyse. Car si le taux de piratage de son top 50 à partir de notre échantillon se révèle peu élevé, en revanche il faut considérer aujourd'hui qu'il est possible de trouver presque chaque titre disponible en offre illégale, avec une recherche plus approfondie sur Internet.

Top 20 des éditeurs ayant le plus de titres piratés (livres)

Ce tableau n'indique pas quels sont les éditeurs les plus piratés en volume (hors poches²⁹), mais quels sont ceux qui ont le plus de titres piratés.

1	Eyrolles	11	Micro Application
2	Gallimard	12	Marabout
3	Dunod	13	Minerva
4	Albin Michel	14	Fayard
5	Hachette ^a	15	Solar
6	Bragelonne ^b	16	Pygmalion
7	Seuil	17	Grasset
8	Flammarion	18	Editions Rustica
9	Larousse	19	Michel Lafon
10	Editions Générales First	20	Plon

^a dont Hachette Pratique et Hachette Littérature / ^b dont Milady

Le classement reflète une double réalité :

- Les gros éditeurs sont aussi les plus piratés.
- Certaines catégories d'édition sont très présentes : c'est le cas des **STM** (Eyrolles, Dunod, First, Micro Application). Les éditeurs de STM disposent, il faut le noter, de catalogues de livres numériques particulièrement bien fournis. Une bonne partie des fichiers piratés semblent l'être, d'après notre analyse, depuis des plateformes légales de téléchargement.
- La présence de Bragelonne dans le top 10 témoigne de la forte présence d'ouvrages de fantastique et science-fiction au sein des fichiers piratés.

²⁸ Les titres concernés sont notamment, pour les romans : *Millenium* de Stieg Larsson, *Le Symbole Perdu* de Dan Brown, *Le Miroir de Cassandre* de Bernard Werber. Pour la jeunesse : *Twilight* de Stephene Meyer et *Les Chevaliers d'Emeraude* d'Anne Robillard. Pour la catégorie pratique : *La Méthode Dukan illustrée* de Pierre Dukan.

²⁹ Le classement ne tient pas compte des rééditions au format poche. Les éditeurs les plus piratés dans ce domaine sont, dans l'ordre : Le Livre de Poche, J'ai Lu, Pocket, J'ai Lu et Fleuve Noir.

Top 12 des éditeurs ayant le plus de titres piratés (BD)

1	Delcourt
2	Soleil
3	Dargaud
4	Glénat
5	Casterman
6	Dupuis
7	Panini Comics
8	Le Lombard
9	Vents d'Ouest
10	Les Humanoïdes Associés
11	Futuropolis
12	Ki-oon

Le classement est constitué des gros éditeurs francophones auxquels s'ajoutent deux éditeurs plus « petits » en termes de chiffre d'affaires : Futuropolis (détenu par Gallimard et Soleil) et Ki-oon.

Il est à comparer avec les parts de marché des éditeurs :

1	Glénat	13 %
2	Kana	9 %
3	Delcourt	8,3 %
4	Dargaud	8,2 %
5	Dupuis	8 %
6	Soleil	6,8 %
7	Casterman	5,5 %
8	Pika	4,7 %
9	Le Lombard	3,8 %
10	Bamboo	3,1 %
11	Kurokawa	2,8 %
12	Panini	2,8 %

Parts des ventes en exemplaires. Source : ACBD, 2010³⁰.

Nous rappelons que, comme pour « EbookZ 1 », nous n'avons analysé que les éditeurs dont les titres sont disponibles en téléchargement illégal, ce qui exclut le streaming. Celui-ci est particulièrement répandu pour les mangas, dont les éditeurs sont de fait parmi les plus piratés.

³⁰ http://www.acbd.fr/images/stories/acbd_bilan_2010.pdf, Gilles Ratier, secrétaire général de l'ACBD (Association des critiques et journalistes de bande dessinée).

Focus — La bande dessinée pirate sur le réseau ed2K

Nous avons analysé l'ensemble des fichiers de BD pirates disponibles sur le réseau ed2K, grâce au moteur de recherche du principal agrégateur de liens de téléchargement, qui comporte plus de 25 000 références (presque totalement en français).

Top 20 des éditeurs de BD ayant le plus de titres piratés sur ed2K

	Editeur	Titres	% du total
1	Dupuis	1863	8,2 %
2	Dargaud	1630	7,2 %
3	Semic	1605	7 %
4	Glénat	1499	6,6 %
5	Marvel France	1387	6,1 %
6	Casterman	1226	5,4 %
7	Delcourt	1180	5,2 %
8	LUG	945	4,2 %
9	Soleil Productions	922	4 %
10	Panini Comics	759	3,3 %

Top 20 des scénaristes de BD ayant le plus de titres piratés sur ed2K

	Scénariste	Titres	% du total
1	Cauvin, Raoul	294	1 %
2	Charlier, Jean-Michel	191	0,7 %
3	Greg	183	0,6 %
4	Duchâteau, André-Paul	182	0,6 %
5	Dufaux, Jean	162	0,6 %
6	Desberg, Stephen	154	0,5 %
7	Cothias, Patrick	146	0,5 %
8	Corbeyran, Éric	141	0,5 %
9	Toriyama, Akira	132	0,5 %
10	Lee, Stan	124	0,4 %

Top 20 des dessinateurs de BD ayant le plus de titres piratés sur ed2K

	Dessinateur	Titres	% du total
1	Tibet	137	0,5 %
2	Toriyama, Akira	132	0,5 %
3	Vandersteen, Willy	99	0,3 %
4	Buscema, John	96	0,3 %
5	Hermann	92	0,3 %
6	Morris	88	0,3 %
7	Hojo, Tsukasa	86	0,3 %
8	Kirby, Jack	84	0,3 %
9	Clamp	81	0,3 %
10	Graton, Jean	80	0,3 %

Année de publication

Un même ouvrage, piraté dans une réédition des années 1990, peut avoir été publié originellement chez un autre éditeur dix ans plus tôt. Les chiffres qui suivent ont donc une valeur indicative. Ils tiennent compte, dans la mesure du possible, de la première date de publication et non pas des rééditions ultérieures.

	1900- 1950	1951- 1970	1971- 1990	1991- 2000	2001- 2006	2007- 2008	2009	2010
Livres	0,5 %	0,9 %	4,3 %	12,2 %	36,1 %	18,7 %	14,1 %	12,6 %
BD	0,0 %	0,0 %	4,5 %	10,7 %	28,8 %	22,6 %	13,6 %	18,5 %
Total	0,4 %	0,7 %	4,3 %	11,8 %	34,4 %	19,6 %	14,0 %	14,0 %

Cette année encore, plus de **2 ouvrages piratés sur 3, livres comme BD, datent de moins de 10 ans**. 26,7 % des livres et 32,1 % des BD ont été publiés il y a moins de 2 ans (26,1 % et 33,1 % publiés il y a moins de 4 ans dans « EbookZ 1 »). Le piratage semble toucher des productions de plus en plus récentes, et pas seulement un fonds de catalogue probablement épuisé.

Toutefois, contrairement aux films et à la musique, le piratage des livres n'est pas encore immédiat : il touche encore minoritairement les nouveautés en rayon. Ainsi le piratage concerne pour plus de la moitié la période allant de 2001 à 2008, des titres « long-sellers ». En matière de livres pirates, les internautes semblent chercher l'ouvrage de valeur sûre ou répondant à un besoin (ou une curiosité) immédiat, que la nouveauté à proprement parler.

Origine des ouvrages

Rappelons que nous n'étudions que les œuvres éditées par des éditeurs français, en français (traduites ou non), et que nous excluons les traductions d'œuvres françaises à l'étranger.

	France	Etranger (traductions)
Livres	61 %	39 %
BD	84 %	16 %

Le tableau est à comparer avec la part des traductions dans la production éditoriale : 14,3 % selon les chiffres de *Livres-Hebdo*/Electre en 2009. Par ailleurs 99 % des ouvrages traduits et piratés proviennent des versions officielles parues chez les éditeurs autorisés. Seules quelques rares exceptions (certaines versions des romans de J.K. Rowling par exemple) sont le fruit de traductions officieuses réalisées par des fans, généralement impatients de diffuser les textes avant leur sortie en France.

Catégories (livres)

1	11-ROMANS SF ET FANTASTIQUES	21,8 %
2	21-PRATIQUE ET LOISIRS	17,3 %
3	8-ROMANS	11,1 %
4	9-ROMANS POLICIERS	10,7 %
5	3-INFORMATIQUE	10,2 %
6	2-SCIENCES TECHNIQUES ET MEDECINE	6 %
7	4-SCIENCES HUMAINES ET SOCIALES	5,2 %
8	16-JEUNESSE	5,1 %
9	1-ENSEIGNEMENT	3,3 %
10	13-DOCUMENTS ET ESSAIS	3 %
11	6-RELIGION ET ESOTERISME	1,6 %
12	14-HUMOUR	1,4 %
13	5-PHILOSOPHIE	0,9 %
14	7-DICTIONNAIRES ET ENCYCLOPEDIES	0,9 %
15	12-THEATRE ET POESIE	0,7 %
16	10-ROMANS EROTIQUES	0,5 %
17	15-MEMOIRES ET BIOGRAPHIES	0,4 %
18	20-BEAUX ARTS	0,4 %

On peut regrouper ces sous-catégories en 5 catégories principales, d'après le découpage de *Livres-Hebdo* utilisé dans les autres études d'ElabZ :

- 1 – LITTÉRATURE : 44,8 %
- 2 – PRATIQUE (incluant STM, Enseignement, Informatique, Dictionnaires) : 37,7 %
- 3 – ESSAIS (SHS, Documents, Religion et Esotérisme, Humour, Philosophie, Mémoires) : 12,5 %
- 4 – JEUNESSE : 5,1 %
- 5 – BEAUX LIVRES : 0,4 %

Observations :

- **Une catégorie dominante : les romans SF et fantastiques**, avec en tête des best-sellers internationaux d'auteurs comme J.K. Rowling, J.R.R. Tolkien, Stephenie Meyer ou Bernard Werber. La présence d'une nouvelle *team** pirate très active spécialisée dans cette littérature (cf. p. 6) explique notamment pourquoi cette catégorie se retrouve en première position.
- **Une forte présence d'ouvrages pratiques et de STM**, qui partagent une même caractéristique : ils sont facilement consultables sur des supports numériques, car ils ne demandent pas une lecture linéaire et répondent à un besoin immédiat (scolaire, professionnel ou personnel) en termes d'information. Au sein de cette catégorie, les ouvrages Pratique & Loisirs et Informatique arrivent en tête avec respectivement 17,3 % et 10,2 % de notre échantillon.
- **Comparaison avec les ventes d'ouvrages papier**³¹ : si les ouvrages pratiques apparaissent dans des proportions similaires sur le marché légal comme illégal, en revanche la littérature est sur-représentée (44,8 % contre 24,5 % en nombre d'exemplaires vendus), notamment à cause de la forte présence de romans fantastiques et de science-fiction. Les ouvrages de beaux-arts sont encore plus minoritaires, puisque leur valeur tient d'abord à leur fonction de

³¹ « Repères statistiques France 2010 », SNE.

« livre-objet », dont la forme compte autant que le contenu et rend le piratage moins aisé et moins pertinent. Enfin, la différence principale tient aux ouvrages Jeunesse dont la part est presque quatre fois inférieure en offre illégale, ce qui peut s'expliquer par la même raison que pour les beaux-livres, mais également par le fait que les lecteurs à qui s'adressent ces ouvrages sont moins susceptibles de les pirater.

	% du chiffre d'affaires	% des exemplaires vendus
1 - LITTÉRATURE	23,6	24,5
2 - PRATIQUE	35,9	33
3 - ESSAIS	11,8	7
4 - JEUNESSE	14,6	21,2
5 - BEAUX-ARTS	3,5	1,5

Focus — L'édition professionnelle

Dans une acception large, l'édition professionnelle, notamment de STM, recouvre dans notre échantillon une part importante de titres destinés au grand public (par exemple des ouvrages de vulgarisation scientifique) ou aux étudiants (notamment en informatique).

Si l'on entend l'édition professionnelle au sens restreint, c'est-à-dire des publications destinées au monde professionnel (par exemple : fiscalité d'entreprise, ouvrages d'ingénierie, etc.), on remarque qu'il existe plusieurs communautés d'échanges pirates spécialisées, aux États-Unis (venant principalement de gros éditeurs comme Elsevier ou John Wiley & Sons), avec des catalogues pirates de milliers de fichiers disponibles. En France, selon nos constatations, le piratage de ce secteur est plus occulte : à l'exception d'ouvrages informatiques et techniques présents dans nos catégories 3 et 2 (rangs 5 et 6 du tableau page 29), ainsi que de quelques ouvrages de référence juridiques ou financiers, il passe encore beaucoup par des FTP privés ou des envois par email au sein de l'entreprise. Il est donc évident que l'édition professionnelle est sous-évaluée dans notre échantillon par rapport à la réalité du piratage.

En ce qui concerne la BD, la répartition des fichiers de notre échantillon est la suivante :

Albums	85 %
Comics	10,1 %
Manga	4,5 %

Les albums représentent donc la très grande majorité des BD de notre échantillon. Cela ne signifie pas que les comics et les mangas soient absents du piratage en ligne : le faible nombre de comics s'explique par le pourcentage massif de BD en version originale (non inclus dans notre échantillon) ; le faible nombre de mangas provient du fait qu'ils sont surtout accessibles en « streaming », c'est-à-dire en feuilletage en ligne sur des sites dits de scantrad* (cf. « EbookZ 1 »).

Prix moyen

Le prix de l'offre légale de livres, papier comme numérique, est l'un des éléments déterminants pour comprendre le piratage. Le tableau ci-dessous indique le prix en offre légale papier des ouvrages piratés (il tient seulement compte des ouvrages encore disponibles neufs à la vente³²).

	< 5€	5 à 10€	11 à 20€	21 à 30€	> 30€	Moyenne
Livres	4,7 %	39,1 %	24,6 %	13,0 %	8,9 %	14,4 €
Livres hors poches	1,8 %	17,1 %	35,6 %	19,5 %	13,3 %	18,5 €
BD	0,8 %	28,4 %	54,7 %	4,5 %	3,3 %	12,8 €

Le prix moyen du livre papier en vente légale des ouvrages piratés, toutes catégories confondues, s'élève à **14,4 € pour les livres** (18,5 € hors poches) et **12,8 € pour les bandes dessinées**.

Ces prix sont à rapprocher des prix de vente moyens en offre légale constatés par l'étude Edistat de septembre 2010 : « Le prix moyen se situe à 12,50 euros pour un livre électronique, soit une décote de 25 % par rapport à la version papier (16,57 euros) ».

Prix moyen par catégorie :

20-BEAUX ARTS	29 €
3-INFORMATIQUE	28,7 €
1-ENSEIGNEMENT	28,3 €
2-SCIENCES TECHNIQUES ET MEDECINE	27,3 €
4-SCIENCES HUMAINES ET SOCIALES	22,8 €
21-PRATIQUE ET LOISIRS	18,2 €
18-BD COMICS	17,7 €
7-DICTIONNAIRES ET ENCYCLOPEDIES	15,5 €
17-BD ALBUMS	14 €
5-PHILOSOPHIE	13,9 €
6-RELIGION ET ESOTERISME	13,7 €
15-MEMOIRES ET BIOGRAPHIES	13,7 €
16-JEUNESSE	13,1 €
13-DOCUMENTS ET ESSAIS	11,4 €
11-ROMANS SF ET FANTASTIQUES	10,3 €
8-ROMANS	10,1 €
9-ROMANS POLICIERS	9,8 €
14-HUMOUR	9,1 €
12-THEATRE ET POESIE	7 €
19-BD MANGA	6,9 €
10-ROMANS EROTIQUES	4 €

³² Dans de nombreux cas, les ouvrages sont disponibles à la vente parallèlement en édition grand format et poche. Nous n'avons dans ce cas retenu que le prix au format poche, considérant que c'était à ce prix, le plus bas possible, qu'il fallait comparer l'offre illégale.

Prix moyen par catégorie hors poches (livres) :

20-BEAUX ARTS	29,0 €
3-INFORMATIQUE	28,7 €
1-ENSEIGNEMENT	28,2 €
2-SCIENCES TECHNIQUES ET MEDECINE	27,3 €
4-SCIENCES HUMAINES ET SOCIALES	27,2 €
6-RELIGION ET ESOTERISME	19,8 €
11-ROMANS SF ET FANTASTIQUES	19,5 €
21-PRATIQUE ET LOISIRS	18,2 €
15-MEMOIRES ET BIOGRAPHIES	18,0 €
7-DICTIONNAIRES ET ENCYCLOPEDIES	17,6 €
16-JEUNESSE	17,5 €
8-ROMANS	16,2 €
13-DOCUMENTS ET ESSAIS	15,2 €
5-PHILOSOPHIE	15,0 €
9-ROMANS POLICIERS	14,0 €
14-HUMOUR	10,8 €
12-THEATRE ET POESIE	6,5 €

Le prix du livre papier, s'il est peu influent concernant une grande partie des ouvrages de fiction achetables au format poche pour moins de 10 euros, est-il un facteur important de piratage d'ouvrages pratiques et STM, dont la moyenne cumulée s'élève à 22,8 € euros ? La question reste ouverte.

Disponibilité des ouvrages

Il est enfin intéressant de faire ressortir le **nombre d'ouvrages piratés non disponibles à la vente papier** :

Livres	9,7 %
BD	8,2 %

Ces taux ont beaucoup évolué depuis l'étude « EbookZ 1 » puisqu'ils étaient 3 à 4 fois supérieurs. On peut l'expliquer par le développement des usages de lecture numérique (avec le développement des liseuses et autres appareils nomades) qui crée une demande pour des livres plus récents.

Synthèse des résultats

- **Une offre de livres numériques pirates encore faible (moins de 2 % de l'offre légale papier) mais en progression.**

Principales évolutions entre « EbookZ 1 » et « EbookZ 2 » :

- **Accroissement global de l'offre illégale pour les ebooks**, toutes catégories de livres confondues, avec environ 2 000 à 3 000 titres disponibles.
- En ce qui concerne la BD, près de 30 000 titres existent (hors streaming) mais seulement **6 000 à 7 000** sont disponibles et facilement accessibles.
- Evolution des usages : **forte baisse des échanges en peer to peer au profit du téléchargement direct** qui devient la pratique dominante en matière de piratage de livres, avec **près de 3 livres sur 4** piratés de cette manière.
- **Un piratage qui touche toujours les best-sellers mais aussi de plus en plus les nouveautés**, au détriment du fonds de catalogue de titres indisponibles à la vente.
- **Apparition de fichiers illégaux au format ePub** et développement des packs multiformats de haute qualité.
- **La taille moyenne des fichiers est en baisse (20,1 Mo pour les livres** contre 29,7 Mo l'année dernière), ce qui est notamment dû à la multiplication de fichiers au format texte (ePub et PDF), parfois directement piratés depuis l'offre légale.

Tout se passe comme si la période de transition se prolongeait concomitamment au décollage du marché du livre numérique en France. Les recommandations faites l'année dernière aux éditeurs et aux acteurs de la chaîne du livre, en conclusion d'« EbookZ 1 », restent en cela pertinentes.

Le piratage n'est pas encore massif mais il s'accélère, et le délai de piratage des nouveautés semble plus court. A titre d'exemple, à l'heure où nous écrivons, le best-seller de Stéphane Hessel, *Indignez-vous !*, est disponible sur les réseaux pirates seulement quelques semaines après le début de son « buzz » médiatique.

Toutefois, si l'offre se développe, cette multiplication des références de fichiers pirates ne signifie pas nécessairement que ceux-ci soient réellement actifs : en peer to peer par exemple, beaucoup d'entre eux ne sont partagés que par peu de personnes et donc difficilement téléchargeables.

Le meilleur moyen d'enrayer le développement du piratage est la mise en place d'une offre légale attractive et de qualité. Propos aujourd'hui partout relayés par les acteurs du livre dans les nombreux débats professionnels sur le numérique, y compris par les élus au sein de l'Assemblée nationale.

Reste à déterminer précisément ce que constitue l'attractivité de cette offre et dans quelle mesure le marché rencontre les usages (et réciproquement), car il existe un risque d'accroissement du piratage avec le développement du légal numérique, tant que la qualité du légal ne l'emporte pas sur la qualité des fichiers pirates.

Lexique

binaire	fichier qui ne peut être lu au format texte (par exemple : un format image). Les fichiers échangés dans les newsgroups le sont sous cette forme, découpés en plusieurs parties qui seront réassemblées par la suite par les téléchargeurs.
direct download (DDL)	fichier disponible en téléchargement direct depuis un lien vers un site de stockage spécialisé (les plus connus étant MegaUpload et RapidShare).
darknet	réseau privé de partage de fichiers, dont l'accès est restreint à certains utilisateurs.
DRM	acronyme de « Digital Rights Management », qui recouvre les procédés de protection technique des droits pour les contenus numériques. Les éditeurs de livres français n'ont pas signé les accords Olivennes et se réservent le droit d'utiliser des DRM.
ebook	traduit par « livre électronique » ou « livrel », il s'agit de la version numérique d'un livre, consultable depuis un ordinateur, un appareil portable ou un lecteur spécialisé appelé <i>reader</i> .
eDonkey (ed2k)	système de partage de fichiers en <i>peer to peer</i> , dont le logiciel le plus populaire est eMule.
fake	fichier dont le nom ne correspond pas à son contenu. Les <i>fakes</i> sont créés soit par des personnes souhaitant profiter de la notoriété de certains contenus pour diffuser le leur, soit par les producteurs eux-mêmes afin de polluer le téléchargement illégal.
leecher	(en anglais : « sangsue ») utilisateur de <i>peer to peer</i> qui télécharge sans partager ses fichiers.
peer to peer (P2P)	abrégié en P2P, il s'agit d'un système décentralisé d'échanges de contenus numériques, d'ordinateur à ordinateur (par opposition au modèle client-serveur).
reader	traduit par « liseuse », il s'agit d'un appareil de lecture d'ebooks. Les leaders du marché sont aujourd'hui Sony (Sony Reader) et Amazon (Kindle). On appelle également ainsi les applications pour ordinateur ou appareils portables comme l'iPhone, offrant une fonction similaire (comme Stanza ou eReader).
release	diffusion sur les réseaux d'un fichier numérique, en général illégal. On parle de « release group » à propos de l'équipe qui effectue cette diffusion.
scantrad (scanlation)	manga numérisé, traduit (généralement depuis l'anglais ou le japonais), et diffusé illégalement et la plupart du temps gratuitement sur Internet. Il s'agit en général de mangas non disponibles à la vente en France.
seeder / seed	(en anglais : « semeur ») utilisateur de <i>peer to peer</i> qui partage du contenu (des copies complètes de fichiers), par opposition au <i>leecher</i> . Un « seed » est un utilisateur partageant un fichier complet.
team	équipe de création et de diffusion de fichiers pirates, parfois structurée de façon quasi professionnelle.
téléchargeur	personne qui télécharge des fichiers sur Internet.
torrent	désigne soit le protocole d'échanges en <i>peer to peer</i> BitTorrent, soit le fichier de métadonnées contenant les informations sur le fichier à télécharger, soit encore le fichier lui-même.
tracker	désigne à la fois le fichier de référence d'un contenu diffusé en torrent ainsi que les sites indexant ces fichiers.
warez	tout contenu protégé et diffusé illégalement sur les réseaux.

Il est interdit de reproduire cette étude partiellement ou intégralement, sur quelque support et pour quelque usage que ce soit, sauf ceux qui sont définis par la loi, sans l'autorisation préalable et écrite du MOTif.